

口頭能力試験開発についての中間報告

－ 日本語能力試験の現在の取り組み －

青山 眞子

はじめに

日本語能力試験（以下、「能力試験」と略す）は、日本国内及び国外において日本語を母語としない者の日本語能力を測定し、認定することを目的としている。1984年から毎年1回、12月上旬の日曜日に実施され、2002年で19回目を迎えた。能力試験は困難度の最も高い1級から最も低い4級まで4つの試験があり、それぞれの級は「文字・語彙」「聴解」「読解・文法」の3つの類によって構成されている。それぞれの級がどのような日本語能力を要求しているかは「認定基準」に記されている⁽¹⁾。

『日本語能力試験 結果の概要 2001』によると、ほぼ10年前の1990年度は、22の国（50都市）で実施されていたが、2001年度は国内6地域、国外38の国（89都市）で実施され、実施地数はこの10年間に大きく増加している。また、受験者数は1990年度、国内18,167人、国外24,620人、全体で4万人程度であったが、2001年度は、国内49,094人、国外178,499人、全体で22万人を超えるまでに至っている。

第1回目の能力試験を実施した1984年当時は、日本語学習者の日本語能力を測定する試験は能力試験が唯一であったが、その後、1996年にはジェトロビジネス日本語能力テスト⁽²⁾、1998年には日本語力測定試験⁽³⁾、2002年には日本留学試験⁽⁴⁾が実施され、試験を取り巻く環境は大きく変化している。

大坪（1995）は、1995年度の時点で、能力試験は「初期のころ抱えていたさまざまな問題を徐々に解決しながらある程度の安定性を示しはじめている」としながらも、いくつかの問題をあげ、その中で「会話・作文試験の実施」の必要性を述べている。1998年、日本語能力試験企画小委員会口頭能力試験部会⁽⁵⁾（以下、「口頭能力試験部会」）が発足し、口頭能力試験の開発に取り組んでいる⁽⁶⁾。

1. 口頭能力試験開発の経緯

能力試験開発当初、「日本語能力試行試験」を5年間実施し、試験実施の可能性に関する調査・研究を行っている。その結果をまとめた『外国人のための日本語能力認定試験に関する調査・研究の経過報告V』によると「試行試験の1回目には面接試験が含まれていたが、その公平、敏速かつ経済的な採点が非常に困難であり、2回目以降は中止することにし、日本語能力

試験でも会話の試験は課さないことにした」とある。

能力試験の受験者が22万人を超える現在、能力試験の一環として口頭能力試験を実施した場合、何らかの方法で限定するとしても、口頭能力試験受験希望者はかなりの数にのぼると思われる。大坪（1995）で述べられているように、口頭能力試験を実施するためには、受験者の発話標本を「公平、敏速かつ経済的に採点」するための評価方法を開発する必要がある。そのうえ、国際交流基金が試験を実施する場合、海外での試験実施が可能でなければならない。このような事情から、能力試験の一環として口頭能力試験を開発する場合、以下の5つの条件を満たす必要がある。

- (1) 多数の受験者を同時に測定できること。
- (2) 発話標本採取における公平性が維持できること。
- (3) 反復受験時における客観性・信頼性を具備すること。
- (4) 実施および評価手順が簡便であるという実用性を具備すること。
- (5) 受験者の学習背景を問わないこと。

上記の条件を満たすような口頭能力試験を考えた場合、インタビューが海外に出向いてインタビューを実施するのは現実的とは言えない。そこで、口頭能力試験部会では機械によって問題を出題する口頭能力試験の開発に取り組むこととなった。

口頭能力試験部会は、1998年11月からほぼ毎月1回会議を開催し、2002年11月には40回目を迎えた。その間、2000年には第1回試行試験を、2001年には第2回試行試験を実施した。第1回試行試験については庄司他（2000）で、第2回試行試験については青山（2002）で一部報告されている。2002年9月現在、部会は伊東祐郎氏（東京外国語大学）、迫田久美子氏（広島大学大学院）、庄司恵雄氏（お茶の水女子大学）、野口裕之氏（名古屋大学大学院）、春原憲一郎氏（海外技術者研修協会）の5名の委員で構成されている⁽⁷⁾。以下、2回の試行試験の実施方法と分析結果について報告する。

2. 第1回試行試験

2.1 実施方法

第1回試行試験は、以下の2つの目的を検証するために実施された。

- (1) インタビューによる出題と、機械（ビデオテープ）（以下、「機械」）による出題で被験者の発話標本にどのような差があるか。
- (2) 機械による出題は可能なのか。

試行試験は、東京と大阪の、合計2カ所の日本語学校で、2000年2～3月に実施した。被験者は、日本の大学に入学することを目的に予備教育課程で学習している計30名の学生で、中国、台湾、韓国など漢字圏が多い。試験は、ウォームアップにつづき、①「音読問題（問題を

読む)」、②「Q&A (質問に答える)」、③「絵説明 (ストーリーを組み立てる)」、④「グラフ説明 (グラフについて説明する)」の4つの問題で構成されている。

被験者はそれぞれ個別の部屋に入り、同一日にインタビューによる出題と、機械による出題の2つの異なる方式による、各30分程度の試験を連続して受けた。試験終了後、試験に関するアンケート調査に回答した。アンケート調査は、問題内容、インタビュー・機械による出題について質問している。

2.2 分析方法と結果

上記の目的(1)を検証するため、インタビューによる出題と、機械による出題の両方式によって得られた被験者の発話標本を量的側面から分析した。具体的には、両方式の文節の数と活用語の異なり語数を、対応のある標本の平均値の差の検定を用いて分析した。その結果、いずれの場合についても方式間で有意差は認められず、両方式の文節数と活用語数という発話標本の量的側面には差がないということが判明した。

試験実施前は、解答するのに十分な日本語の口頭能力があるにも関わらず、機械による出題に緊張や戸惑いを感じ、発話できなくなる被験者がいるのではないかと予測したが、実際にはそのような被験者は見られなかった。また、試行試験終了後のアンケート結果ではインタビューによる出題の方が、機械による出題よりも話しやすいと回答した被験者が多かったが、両方式に対する回答と発話標本の文節数との相関は、人間によるインタビューの方が高かった。つまり、被験者個人の試験方式に対する心理的な抵抗感が発話標本の量に与える影響は、インタビューによる方式よりも、機械による方式の方が小さかった⁽⁸⁾。以上から、上記目的(2)の機械による出題は可能であると判断した。

3. 第2回試行試験の実施

3.1 実施方法

第1回試行試験後、インタビューではなく機械による出題が可能であると判断した。第2回試行試験は、以下の2つの目的で試行試験を実施した。

- (1) 海外において、機械の中でもコンピュータによって出題される試験が実施可能か。国内・外の受験者に受け入れられるのかを検証する。
- (2) この試験で測定しようとする能力、つまり試験の構成概念は何かを明確にする。

試験は2001年8月から10月にかけて、関西国際センター、日本語国際センター、日本語学校(東京)、日本語学校(大阪)、大学(京都)の国内5ヵ所、シドニー日本語センター、バンコク日本語センター、クアラルンプール日本語センター、ジャカルタ日本語センターの海外4ヵ所の合計9ヵ所で実施した。被験者は国内130名、海外76名、合計206名である。国

籍は、中国、マレーシア、インドネシア、タイ、オーストラリアを始め、35ヶ国と多様であり、206名中71名が漢字圏出身者である。

問題は、①「Q & A（質問に答える）」、②「道順説明（目的地までの行き方を説明する）」、③「商品取替（買ってきた辞書が汚れていたのを換えてもらう）」、④「留守番電話（ピアノのコンサートに友達を誘う）」、⑤「写真説明（写真について説明する）」、⑥「絵説明（4つの絵を見てストーリーを作る）」、⑦「グラフ説明（グラフを見て説明する）」、⑧「議論（仕事について議論する）」、⑨「スピーチ（若者の問題について自由に意見を述べる）」の9つの問題で構成されている。問題の前後には、「ウォームアップ」と「クールダウン」がある。

被験者は、試験当日、まず始めに5分程度のコンピュータ試験についての説明を聞き、そのあと約25分のコンピュータ試験を受ける。実施地によっては、コンピュータ試験のあとに、OPI⁽⁹⁾、Can-do-statements 調査⁽¹⁰⁾を実施した。コンピュータ試験の後、被験者は試験についてのアンケートに回答した。

3.2 分析方法と結果

第2回試行試験は、コンピュータを使用したため、初めの数回はコンピュータ、電源設備などの環境による不備、スタッフの不慣れなどから、出題時に問題が起こったり、評価可能なレベルでの録音できていないということもあった。しかし、試行試験の回数を重ねるに従い、スタッフも慣れ、国内・外の実施において、十分評価可能な発話標本を得ることができ、無事試行試験を終えることができた。試行試験終了後のアンケート調査からも、被験者からコンピュータ試験に対して好意的な評価が得られ、コンピュータ試験が国内・外で実施可能であり、受験者からも概ね受け入れられたと言える。上記の目的(2)については、現在分析中である。

おわりに

近年、試験を取り巻く環境だけではなく、教育関係者の試験に対する意識も徐々にではあるが変化してきている。以前は、試験の結果は絶対的なものであり、試験の信頼性、妥当性を疑うことはほとんどなかった。しかし、現在では、試験実施・作成・分析など、試験開発に対する興味、関心は高まりつつあり、きちんとした手続きに基づいて作成された試験は信頼すべき結果を与えるが、得点は絶対的なものではないという考え方が広がり、浸透しつつある。バックマン・パーマー(2000)の1章に「言語テスト関係者は、ほとんど魔法にも似た手順とやり方で『最良』のテストを作るものだ」という思い込みをしている人がよくいる。」とある。バックマン・パーマーの言うように、「最良」のテストというものはなく、テストにはそれぞれ改善すべき点があって当然である。

しかしながら、能力試験のように、受験者が22万人を超えるという大規模テストの場合、試

験内容とその結果が、受験者と教育現場に与える影響は大きく、利害関係の大きいテストであることは間違いない。国内・外で公平かつ経済的に、発話標本を採点できる口頭能力試験を開発することは簡単なことではないが、口頭能力試験の開設が各方面から待たれていることは言うまでもない。

第2回試行試験の調査結果については現在分析中であるが、試験開発当初から目指してきた、機械の出題による口頭能力試験の実施可能性が2回の試行試験によって検証された。部会発足以来、4年を経過しようとしているこの段階になってはじめて、海外での口頭能力試験実施が可能であるという感触を部会でも得つつある。1998年から2回の調査を経て得られた成果は、報告書にまとめられ、その内容は2002年度末の日本語能力試験企画小委員会、実施委員会で報告される予定である。

〔注〕

- (1) 「認定基準」については、毎年発行されている『日本語能力試験結果の概要』を参照。
- (2) 詳しくは、日本貿易振興会（2001）を参照。
- (3) 詳しくは、日本語学研究所（1998）を参照。
- (4) 詳しくは、財団法人日本国際教育協会（2002）を参照。
- (5) 能力試験には審議を行うため、いくつかの部会が設置されている。まず、重要事項を審議するための、日本語能力試験実施委員会、その下に試験の実施計画の策定及び実施状況の把握、試験問題および試験結果の分析および評価などを主に取り扱う企画小委員会が存在する。口頭能力試験部会は企画小委員会の下部組織である。
- (6) 試験の開発は口頭能力試験部会で取り組んだ。したがって、今回の報告内容は執筆者、青山眞子一人の成果ではなく、部会全体の成果であるが、本稿の文責は青山にあることを申し添える。
- (7) 2002年8月以前は、小出慶一氏（群馬女子大学）、金澤眞智子氏（京都外国語大学）も委員の一員であった。
- (8) 詳しくは、庄司・青山・金澤・野口（2000）を参照。
- (9) OPIは、全米外国語教育協会（The American Council on the Teaching of Foreign Languages）による口頭能力試験（Oral Proficiency Interview）を指す。
- (10) Can-do-statements 調査は、日本語能力を読む、書く、話す、聞くについての61の質問に答えることによって、被験者に自己評価してもらう調査である。

〔参考文献〕

青山眞子（2002）「口頭能力試験の調査・開発について」第15回日本言語テスト学会研究例会
大坪一夫（1995）「日本語能力試験」『日本語教育』86号別冊 pp.72-81

- 国際交流基金・財団法人日本国際教育協会（2001）『「日本語能力試験 出題基準」の平成14（2002）年度からの一部改訂について』 凡人社
- _____（2002）『日本語能力試験 出題基準』 凡人社
- 財団法人日本国際教育協会・国際交流基金（2002）『日本語能力試験 結果の概要 2001 年度版』
- _____（2002）『日本語能力試験の概要 2001 年度版』
- _____（2001）『平成13年度版日本語能力試験 分析評価に関する報告書』
- 財団法人日本国際教育協会（2002）『平成14年度日本留学試験（第1回）試験問題』 桐原書店
- 庄司恵雄・青山眞子・金澤眞智子・伊東祐郎・野口裕之（2000）「大規模口頭能力試験開発に関する基礎的研究—発話標本採取法の検討—」『日本語教育』116号 日本語教育学会 pp.109-118
- 日本語教育学会（1985）『外国人のための日本語能力認定試験に関する調査・研究の経過報告V』
- 日本語学研究所（1998）『日本語力測定試験ガイダンスセット』 日本語学研究所
- 日本語能力試験実施委員会・日本語能力試験企画小委員会（2002）『日本語能力試験分析評価に関する報告書 1990年から1999年度までの分析報告書 10年度分試験結果の分析』 国際交流基金・財団法人日本国際教育協会
- 日本貿易振興会（2001）『ジェトロビジネス日本語能力テストガイド 第2版』 JETRO
- 野口裕之・青山眞子・廣利正代（2001）「日本語能力試験の因子分析的検討」『2001年度日本語教育学会秋季大会予稿集』 日本語教育学会 pp.103-108
- L.F. バックマン・A.S. パーマー（2000）『＜実践＞言語テスト作成法』 大修館書店